

# 多媒體技術與應用

## Spring 2021

Instructor : Yen-Lin Chen(陳彥霖), Ph.D.

Professor

Dept. Computer Science and Information Engineering

National Taipei University of Technology

# Project 4

分類器練習

# 個人專案要求

- 在Python中引入scikit-learn函式庫，利用函式庫中所提供的函式

**datasets.load\_iris()**

- 利用課程中所教的鳶尾花資料集進行練習
- 鳶尾花資料集中，Data為鳶尾花的特徵，包含花萼長度，花萼寬度，花瓣長度，花瓣寬度四種的特徵資料，Target為鳶尾花的屬種*setosa*，*versicolor*，*virginica*三種做為資料的標記。
- 請測試以cascade架構下的Adaboost、KNN二種分類器進行訓練，以各種不同的kernel function、C參數、gamma參數，以及不同的訓練集、測試集比例，觀察調整以上變數的情況下，對於訓練出來的準確率有什麼變化。
- 參考資料：
  - <https://towardsdatascience.com/knn-using-scikit-learn-c6bed765be75>
  - <https://www.itread01.com/content/1547087057.html>

# 小組專案要求

- 在Python中引入scikit-learn函式庫，利用函式庫中所提供的函式 **datasets.load\_wine()**
  - 利用波士頓房價資料集進行練習。
  - 波士頓房價資料集中，Data分為  
**‘alcohol’**, **‘malic\_acid’**, **‘ash’**, **‘alcalinity\_of\_ash’**, **‘magnesium’**,  
**‘total\_phenols’**, **‘flavanoids’**, **‘nonflavanoid\_phenols’**, **‘proanthocyanins’**,  
**‘color\_intensity’**, **‘hue’**, **‘od280/od315\_of\_diluted\_wines’**, **‘proline’**  
等特徵資料，Target為**“琴酒”**, **“雪莉”**, **“貝爾摩德”**三種不同的酒種類。
  - 請測試以各種**不同特徵的輸入**(ex.僅用alcohol與ash作為輸入，請自行將Data資料集內容儲存為新的陣列，並輸入進SVM進行訓練)，不同的**kernel function**、**C參數**、**gamma參數**，以及不同的**訓練集**、**測試集比例**，觀察調整以上變數的情況下，準確率有什麼變化。

# 專案繳交要求

- 專案繳交項目：小組報告(PPT+程式檔案)、個人報告(WORD+程式檔案)。
- 小組報告須包含以下內容(壓縮成: **Project4\_第X組\_小組報告.zip**)：
  - 小組成員名單(於PPT中表示)。
  - 執行程式是否有遇到什麼困難，如何解決(於PPT中表示)。
  - 題目的程式碼(.py)(上傳任一組參數的版本即可)
  - 測試過的變數組合及觀察結果(於ppt中表示)。
- 個人報告需包含以下內容(檔名格式：**Project4\_學號\_姓名.zip**)：
  - 本次課程中個人所學、遇到的困難(於WORD中表示)
  - 成果圖(於WORD中表示)(上傳任一組參數的版本即可)
  - 題目的程式碼(.py)(上傳任一組參數的版本即可)
- 本專案應於Windows環境中執行。
- 本專案繳交期限至2021/04/08(四)13:00

# 分數分配

- 小組報告：
  - 遇到困難及解決方法(20%)
  - 測試變數的完整度(每測試一種組合10分)(30%)
    - 小組是如何決定要挑選這些特徵作為輸入
  - 測試中觀察到的差異(50%)
- 個人報告：
  - 個人所學、遇到的困難(20%)
  - 成果圖(KNN 20% , AdaBoost 20%)
  - 練習心得(KNN 20% , AdaBoost 20%)